

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

1/15

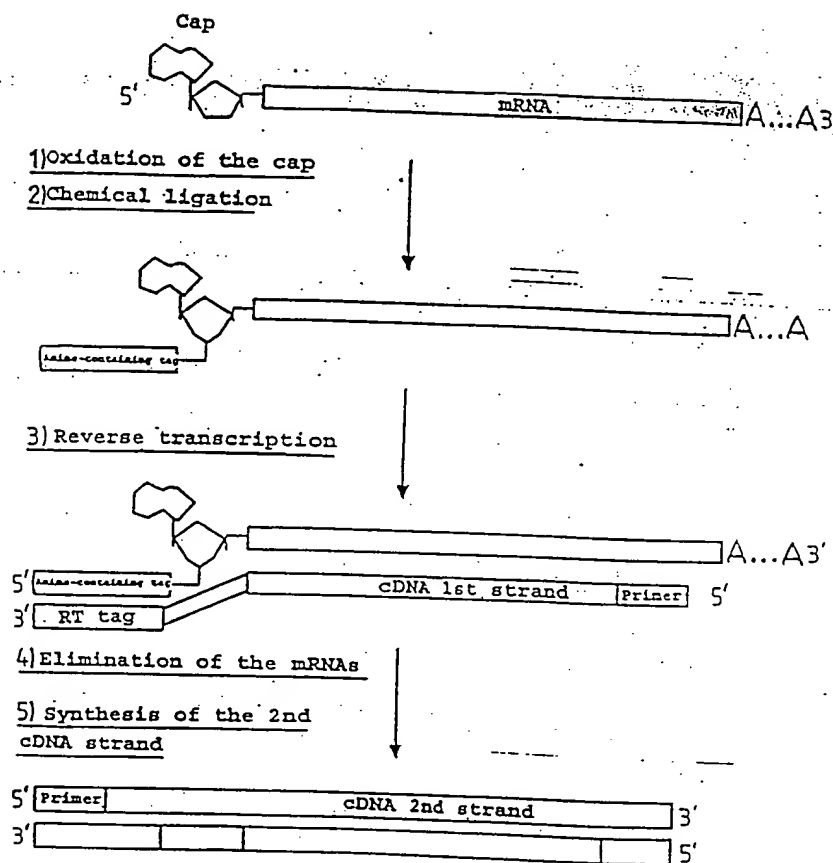


FIGURE 1

2/15

Minimum signal peptide score	false positive rate	false negative rate	proba(0.1)	proba(0.2)
3.5	0,121	0,036	0,467	0,664
4	0,096	0,06	0,519	0,708
4.5	0,078	0,079	0,565	0,745
5	0,062	0,098	0,615	0,782
5.5	0,05	0,127	0,659	0,813
6	0,04	0,163	0,694	0,836
6.5	0,033	0,202	0,725	0,855
7	0,025	0,248	0,763	0,878
7.5	0,021	0,304	0,78	0,889
8	0,015	0,368	0,816	0,909
8.5	0,012	0,418	0,836	0,92
9	0,009	0,512	0,856	0,93
9.5	0,007	0,581	0,863	0,934
10	0,006	0,678	0,835	0,918

FIGURE 2

3/15

Score curves

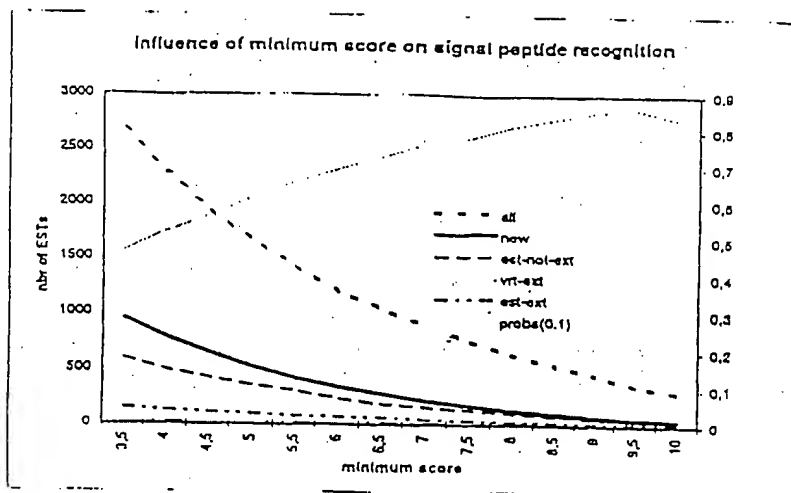


FIGURE 3

4/15

Minimum signal peptide score	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
3,5	2874	947	599	23	150
4	2278	784	489	23	126
4,5	1943	647	425	22	112
5	1657	523	353	21	96
5,5	1417	418	307	19	80
6	1180	340	238	18	68
6,5	1035	280	186	18	60
7	893	219	161	15	48
7,5	753	173	132	12	36
8	636	133	101	11	29
8,5	543	104	83	8	26
9	456	81	63	6	24
9,5	364	57	48	6	18
10	303	47	35	6	15

FIGURE 4

5/15

Tissue	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
Brain	329	131	75	3	24
Cancerous prostate	134	40	37	1	6
Cerebellum	17	9	1	0	6
Colon	21	11	4	0	0
Dystrophic muscle	41	18	8	0	1
Fetal brain	70	37	18	0	1
Fetal kidney	227	116	46	1	19
Fetal liver	13	7	2	0	0
Heart	30	15	7	0	1
Hypertrophic prostate	86	23	22	2	2
Kidney	10	7	3	0	0
Large intestine	21	8	4	0	1
Liver	23	9	6	0	0
Lung	24	12	4	0	1
Lung (cells)	57	38	6	0	4
Lymph ganglia	163	60	23	2	12
Lymphocytes	23	6	4	0	2
Muscle	33	16	6	0	4
Normal prostate	181	61	45	7	11
Ovary	90	57	12	1	2
Pancreas	48	11	6	0	1
Placenta	24	5	1	0	0
Prostate	34	16	4	0	2
Spleen	56	28	10	0	1
Substantia nigra	108	47	27	1	6
Suprarenals	15	3	3	1	0
Testis	131	68	25	1	8
Thyroid	17	8	2	0	2
Umbilical cord	65	17	12	1	3
Uterus	28	15	3	0	2
Non tissue-specific	568	48	177	2	28
Total	2677	947	601	23	150

FIGURE 5

2/15

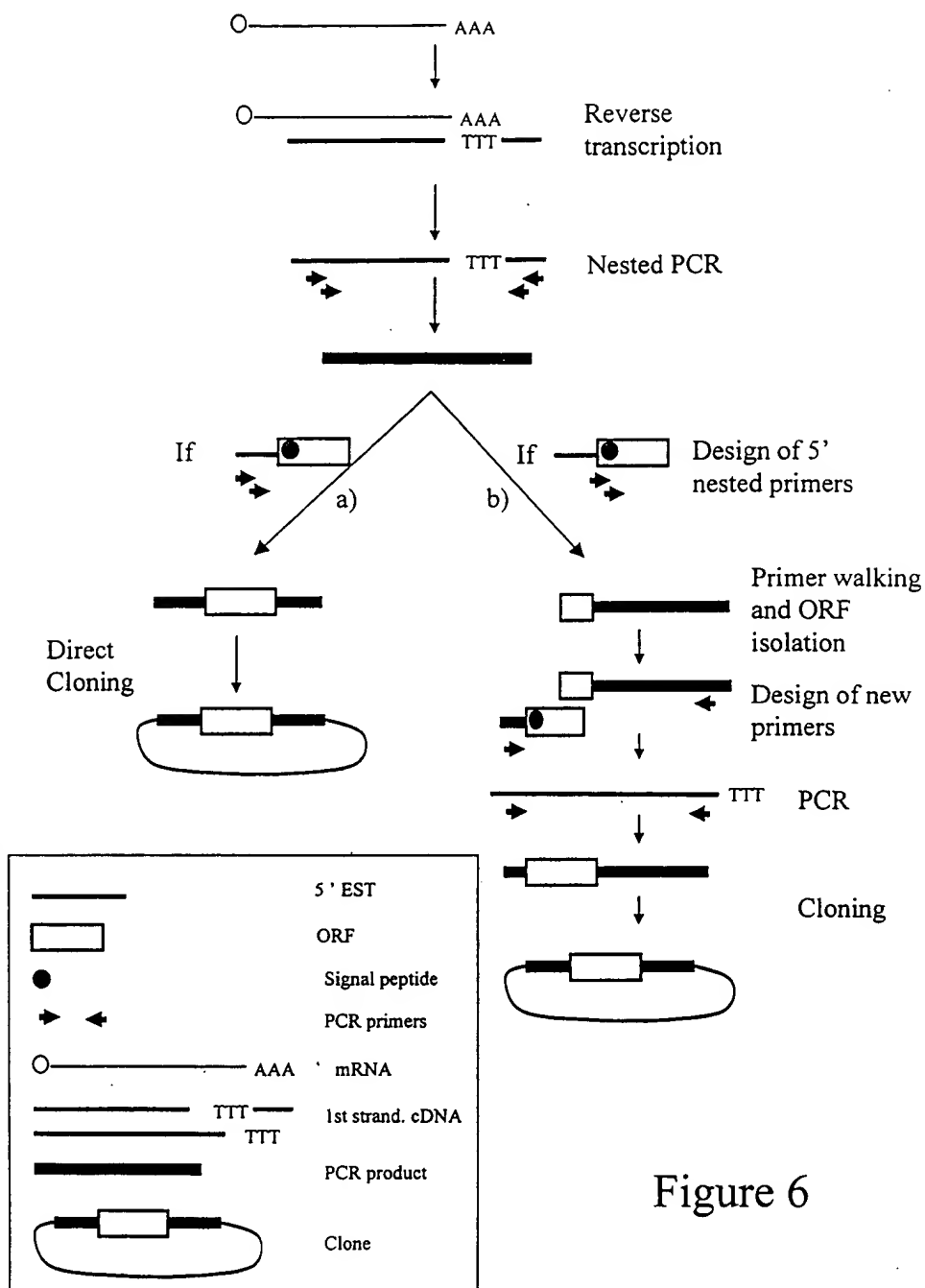
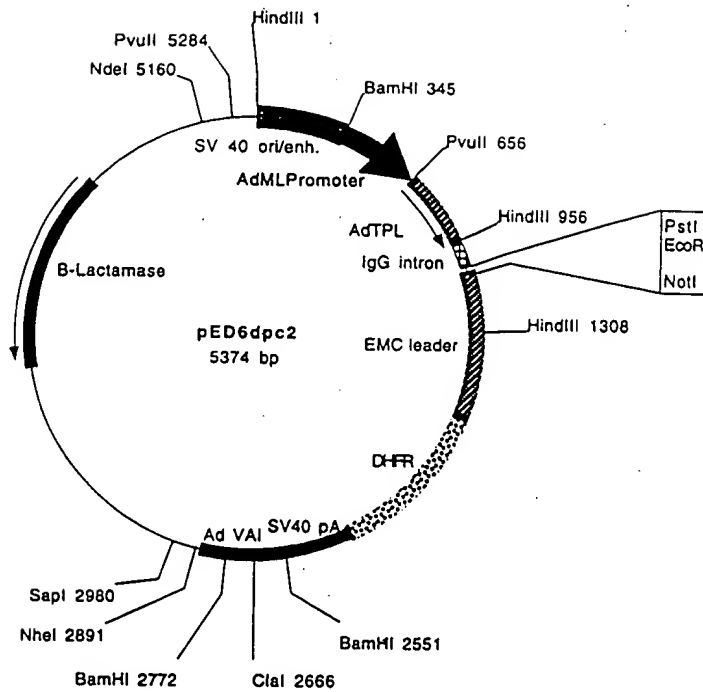


Figure 6

7/15



Plasmid name: pED6dpc2  
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

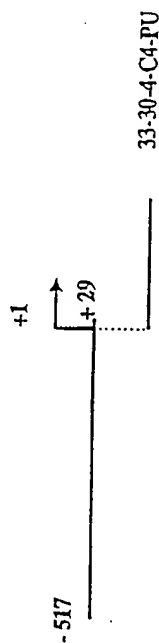
FIGURE 7



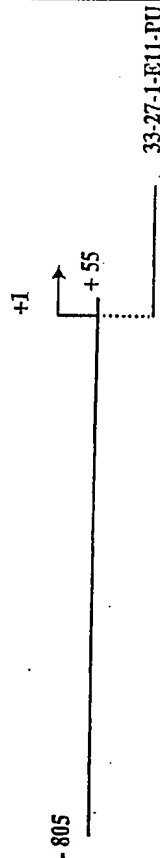
8/15

Description of Promoter structure isolated from SignalTag 5'ESTs

Promoter P13H2



Promoter P15B4



Promoter P29B6

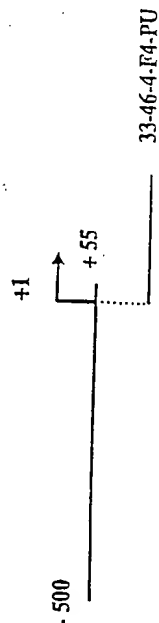


FIGURE 8

9/15

Description of Transcription Factor Binding Sites present on promoters isolated from SignalTag sequences.

Promoter sequence P13H2 (548 bp):

Matrix	Position	Orientation	Score	Length	Sequence
CMYB_01	-502	+	0.963	9	TGTCAGTTG
MYOQ_Q8	-501	-	0.961	10	CCCAACTGAC
S8_01	-444	-	0.960	11	AATAGAAATTAG
S8_01	-425	-	0.966	11	AATCAAAATTAG
DELTAEF1_01	-380	-	0.960	11	GCACACCTCAG
GATA_C	-364	-	0.964	11	AGATAAATCCA
CMYB_01	-348	+	0.958	9	CTTCAGTTG
GATA1_02	-343	+	0.969	14	TTGTAGATAGGACA
GATA_C	-339	+	0.953	11	AGATAGGACAT
TAL1ALPHA47_01	-235	+	0.973	16	CATAACAGATGGTAAG
TAL1BETA47_01	-235	+	0.963	16	CATAACAGATGGTAAG
TAL1BETA172_01	-235	+	0.978	16	CATAACAGATGGTAAG
MYOQ_Q8	-232	-	0.954	10	ACCATCTGTT
GATA1_04	-217	-	0.953	13	TCAAGATAAAGTA
IK1_01	-125	+	0.963	18	AGTTGGGAATTC
IK2_01	-125	+	0.965	12	AGTTGGGAATTC
CREL_01	-123	+	0.962	10	TGGGAATTC
GATA1_02	-96	+	0.960	14	TCAGTGATATGGCA
SRF_02	-41	+	0.961	12	TAAACAAACA
E2F_02	-33	+	0.957	8	TTTAGCGC
MZF1_01	-6	-	0.976	8	TGAGGGGA

Promoter sequence P16B4 (681 bp):

Matrix	Position	Orientation	Score	Length	Sequence
NFY_Q8	-748	-	0.956	11	GGACCAATCAT
MZF1_01	-738	+	0.962	8	CCTGGGA
CMYB_01	-684	+	0.994	9	TGACCGTTG
VMYB_Q2	-682	-	0.965	9	TCCAACGGT
STAT_01	-673	+	0.966	9	TTCTGGAA
STAT_01	-673	-	0.951	9	TTCCAGGA
MZF1_01	-656	-	0.956	8	TTGGGGGA
IK2_01	-451	+	0.965	12	GAATGGGAATTC
MZF1_01	-424	+	0.968	8	AGAGGGGA
SRF_02	-368	+	0.955	12	GAACAAACA
MZF1_01	-216	+	0.960	8	GAAGGGGA
MYOQ_Q8	-190	+	0.961	10	AGCATCTGCC
DELTAEF1_01	-176	+	0.958	11	TCCCACTTCC
S8_01	6	-	0.992	11	GAGGCAATTAT
MZF1_01	16	-	0.966	8	AGAGGGGA

Promoter sequence P29B8 (555 bp):

Matrix	Position	Orientation	Score	Length	Sequence
ARN1_01	-311	+	0.964	16	GGACTCAGTGTGCT
NMYC_01	-309	+	0.965	12	ACTCAGTGTGCTG
USF_01	-309	+	0.965	12	ACTCAGTGTGCTG
USF_01	-309	-	0.965	12	GAAGCAAGTGAAT
NMYC_01	-309	-	0.956	12	GAAGCAAGTGAAT
MYCHAX_Q2	-309	-	0.972	12	CAGCAAGTGAAT
USF_C	-307	+	0.997	8	TCAAGTGC
USF_C	-307	-	0.991	8	GCAAGTGA
MZF1_01	-292	-	0.968	8	CATGGGGA
ELK1_02	-105	+	0.963	14	CTCTCCGAAGCCT
CEB1P64_01	-102	+	0.974	10	TCCGGAAGCC
AP1_Q4	-42	-	0.963	11	AGTGAAGTGAAC
AP1FJ_Q2	-42	-	0.961	11	AGTGAAGTGAAC
PAD6_C	45	+	1.000	9	TGTGGTCTC

Figure 9

10/15

100.0% identity in 125 aa overlap

	10	20	30	40	50	60
SEQ ID NO: 217	MADEE	LEALRRQRLAELQAKHGDPGDAAQQEAKHREAEMRNSILAQVLDQSARARLSNLA				
	X::					
SEQ ID NO: 516	MADEE	LEALRRQRLAELQAKHGDPGDAAQQEAKHREAEMRNSILAQVLDQSARARLSNLA				
	10	20	30	40	50	60
	70	80	90	100	110	120
SEQ ID NO: 217	LVKPEKTKAVENYLIQMARYGQLSEKVSEQGLIEILKKVSQQTEKTTTVKFNRKVMDS					
	::					
SEQ ID NO: 516	LVKPEKTKAVENYLIQMARYGQLSEKVSEQGLIEILKKVSQQTEKTTTVKFNRKVMDS					
	70	80	90	100	110	120

SEQ ID NO: 217 EDDDY  
      :::X  
SEQ ID NO: 516 EDDDY

100.0% identity in 125 aa overlap

FIGURE 10

11/15

CLUSTAL W(1.5) multiple sequence alignment

```

SEQ ID NO: 517      MFCPLKLILLPVLLDYSGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLS
SEQ ID NO: 232      -----MGCVFQSTEDKCIFKIDWTLS
SEQ ID NO: 174      -----MGCVFQSTEDKRIFKIDWTLS
SEQ ID NO: 175      -----MGCVFQSTVDKCIFKIDWTLS
                      ***** ** *****

SEQ ID NO: 517      PGEHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDNLCNDGSLLLQDVQDVE-----
SEQ ID NO: 232      PGEHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEQADQGTYYICEIRL
SEQ ID NO: 174      PGEHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDNLCNDGSLLLQDVQEQADQGTYYICEIRL
SEQ ID NO: 175      PGEHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEQADQGTYYICEIRL
                      *****

SEQ ID NO: 517      -----
SEQ ID NO: 232      KGESQVFKKAVVLHVLPEEPKGTQMLT-----
SEQ ID NO: 174      KGESQVFKKAVVLHVLPEEPKELMVHVGGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEE
SEQ ID NO: 175      KGESQVFKKAVVLHVLPEEPKELMVHVGGGLIQMGCVFQSTEVKHVTKVEWIFSGR--RAK

SEQ ID NO: 517      -----
SEQ ID NO: 232      -----
SEQ ID NO: 174      IVFRYYHKLMSAEYSQSWGHFQNRVNLVGDI FRNDGSIMLQGVRESDGNYTCSIHLGN
SEQ ID NO: 175      VTRRKHHCVREGSG-----

SEQ ID NO: 517      -----
SEQ ID NO: 232      -----
SEQ ID NO: 174      LVFKKTIVLHVSPEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLPLVILIVKKTTC
SEQ ID NO: 175      -----

SEQ ID NO: 517      -----
SEQ ID NO: 232      -----
SEQ ID NO: 174      GNKSSVNSTVLVKNTKKTNP
SEQ ID NO: 175      -----

```

FIGURE 11

12/15

99.6% identity in 225 aa overlap

```

      10      20      30      40      50      60
SEQ ID NO: 515 PTAVQKEEARQDVEALLSRTVTRQILTQKELRVATQEKEGSSGRCMLTLLGLSFILAGLI
                :
SEQ ID NO: 231                LRVATQEKEGSSGRCMLTLLGLSFILAGLI
                        10      20      30

      70      80      90      100     110     120
SEQ ID NO: 515 VGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIAIIDV
                :
SEQ ID NO: 231 VGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIAIIDV
                40      50      60      70      80      90

      130     140     150     160     170     180
SEQ ID NO: 515 PVPSFSDSDPAAIHDFEKGMTAYLDLLGNCYLMPNTSIVMPPKNLVELFGKLASGRY
                :
SEQ ID NO: 231 PVPSFSDSDPAAIHDFEKGMTAYLDLLGICYLMPNTSIVMPPKNLVELFGKLASGRY
                100     110     120     130     140     150

      190     200     210     220     230     240
SEQ ID NO: 515 LPQTYVVREDLVAVEEIRDVSNLGIYIYQLCNRKSFRLRRRDLLGFKRAIDKCWKIR
                :
SEQ ID NO: 231 LPQTYVVREDLVAVEEIRDVSNLGIYIYQLCNRKSFRLRRRDLLGFKRAIDKCWKIR
                160     170     180     190     200     210

      250     260
SEQ ID NO: 515 HFPNEFIVETKICQE
                :
SEQ ID NO: 231 HFPNEFIVETKICQE
                220

```

FIGURE 12



98.5% identity in 194 aa overlap

```

          90      100      110      120      130      140
SEQ ID NO:519 ARNLPPLTDAQKNKLRHLSVVTLAAKVKCI PYAVLLEALALRNVRQLEDLVIEAVYADVL
               :
SEQ ID NO:158 ARNLPPLTEAQKNKLRHLSVVTLAAKVKCI PYAVLLEALALRNVRQLEDLVIEAVYADVL
          60      70      80      90      100      110

          150      160      170      180      190      200
SEQ ID NO:519 RGSLDQRNQRLEVDYSIGRDIQRQDLSAIAQTLQEWCVGCEVVLSGIEEQVSRANQHKEQ
               :
SEQ ID NO:158 RGSLDQRNQRLEVDYSIGRDIQRQDLSAIAQTLQEWCVGCEVVLSGIEEQVSRANQHKEQ
          120      130      140      150      160      170

          210      220      230      240      250      260
SEQ ID NO:519 QLGLKQQIESEVANLKKTIKVT TAAAAAATSQDPEQHLTELREPASGTNQRQPSKKASKG
               :
SEQ ID NO:158 QLGLKQQIESEVANLKKTIKVT TAAAAAATSQDPEQHLTELREPAPGTNQRQPSKKASKG
          180      190      200      210      220      230

          270
SEQ ID NO:519 KGLRGS AKIWSKSN
               :
SEQ ID NO:158 KGLRGS AKIWSKSN
          240      250

```

88.7% identity in 62 aa overlap

```

          10      20      30      40      50      60
SEQ ID NO:519 MSAEVKVTGQNQEQLLLAKSAKGAALATLIHQVLEAPGVYVFGELLDMPNVRELAESDF
               :
SEQ ID NO:158 MSAEVKVTGQNQEQLLLAKSAKGAALATLIHQVLEAPGVYVFGELLDMPNVRELXARNL
          10      20      30      40      50      60

```

```

SEQ ID NO:519 AS
               .X
SEQ ID NO:158 PP

```

FIGURE 14

[illegible]

**FIGURE 15**